Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Electronic Systems

Torstein Tenstad

Deblurring hyperspectral images of structures embedded in turbid media

Master's thesis in Electronics Systems Design and Innovation
Supervisor: Lise Lyngsnes Randeberg
June 2022

NTNU
Norwegian University of
Science and Technology

Torstein Tenstad

Deblurring hyperspectral images of structures embedded in turbid media

Master's thesis in Electronics Systems Design and Innovation
Supervisor: Lise Lyngsnes Randeberg
June 2022

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Electronic Systems

NTNU
Norwegian University of
Science and Technology

Abstract

The scattering properties of light presents a challenge when capturing hyperspectral images (HSI) of structures covered by turbid media. The captured image is blurry and can be modeled as a convolution between a desired, clear image and a point spread function (PSF). Methods for reconstructing the desired image from the captured images are researched, and three are implemented. The first method uses deconvolution with a closed form expression for the PSF to obtain a reconstruction. The second method reconstructs a clear image from two bands of the HSI, exploiting that different wavelengths are blurred in different amounts. The third method attempts to eliminate the scattering contribution in each pixel of the HSI by estimating the absorption coefficient of the embedded structure. This is done using diffusion theory for light transport and inverse modeling. The methods are tested on a dataset consisting of shapes covered by varying amounts of a turbid media. The reconstructed images from each method are evaluated quantitatively and visually. All three methods provide a measurable and noticeable improvement in the image quality, when applied on an image of a structure covered by 3 mm of turbid media. When the thickness of turbid media is increased to 5 mm, all three methods performs poorly.

Sammendrag

Spredning av lys er en utfordring ved hyperspektral avbildning av strukturer i turbide medium: Bildene blir uskarpe, og kan modelleres som en konvulsjon mellom et ideelt, klart bilde og en punkspredefunksjon. Metoder for å rekonstruere det ideelle bildet utforskes, og tre implementeres. Den første metoden bruker dekonvulsjon med en analytisk utledet punkspredefunksjon. Den andre metoden rekonstruerer det ideelle bilde fra to ulike bånd i et hyperspektralt bilde. Metoden baserer seg på at ulike bølgelengder spres ulikt. Den tredje metoden prøver å fjerne spredningsbidraget i hver piksel av det hyperspektralet bildet ved å estimere absorpsjonskoeffisienten. Dette gjøres v.h.a. diffusjonsteori for lys transport og inversmodellering. De tre metodene testes på et datasett med bilder av former dekket av varierende mengder turbid medium. De rekonstruerte bildene fra hver metode evalueres kvantitativt og visuelt. Alle metodene gir en merkbar og målbar økning i bildekvaliteten på bildet av formene dekket av 3 mm turbid medium. Når tykkelsen økes til 5 mm presterte alle metoder dårlig.

Preface

The work constitutes a master's thesis at the Norwegian University of Science and Technology (NTNU), Department of Electronic Systems. The work was carried out during spring of 2022.

I thank my supervisor Lise Lyngsnes Randeberg for excellent guidance and feedback.

Torstein Tenstad
Trondheim, June 9, 2022

Contents

List of Figures

List of Tables

1 Introduction

Hyperspectral images have more spectral channels than the regular 3-channel RGB images we familiar with. The technology is an efficient and non-invasive diagnostic tool in various disciplines. Applications such as astronomy and geotechnical surveying use hyperspectral imaging to capture the spectral characteristics of objects far away in a wide plane. In industrial applications, hyperspectral cameras perform quality control and material inspection at high rates. In medical diagnostics, hyperspectral imaging is a non-invasive diagnostics tool that can deliver results quickly.[16]

In many hyperspectral imaging applications, the object of interest is covered by a turbid medium, such as the atmosphere, in the case of satellite imaging, or tissue in medical applications. Turbid media is characterized by it's tendency to scatter light, changing it's direction in a random manner.

When observing an object through a turbid media, it's scattering properties makes the object blurry. The effect can be seen on a foggy day or on the bathroom mirror after a shower. This blur is a problem in the imaging applications mentioned above because it obstructs the information the systems are designed to capture. Removing or reducing the blurring effect from the captured images improves the imaging system's performance. With this motivation, the goal of this thesis can defined as follows:

> **Goal:** Explore, develop and evaluate methods for processing hyperspectral images, that reduce blurring created by scattering of light in turbid media.

To achieve this goal, a three step research method was used. First, a literature review was performed. This process is detailed in section 3. Second, three methods where developed based on the findings of the literature review. These methods are largely based on existing methods from the literature, but are adapted to serve the specific goal of this thesis. The adaptation and implementation of the three methods, as well as the dataset used, is presented in section 4. Section 5 shows the result of each method when applied on the dataset. Finally, the methods were evaluated. This is detailed in section 6, and the results of the evaluation are discussed in section 7.

Throughout this thesis, the word *structure* will be used for the thing of interest that is covered by a turbid medium. In the dataset used in this thesis, this is a simple shape. In a general context, this could be more complex, like blood vessels embedded in tissue, or terrain features in a sattelite image.

2 Background theory

The following section presents the theory relevant for the researched methods. Section 2.1 describes hyperspectral cameras and basic optical theory. Section 2.2 presents Boltzmann's transport equation and diffusion theory. Section 2.3 describes the notation of signals used by this thesis. Section presents two denoising algorithms. Section 2.5 describes convolution and two algorithms for deconvolution. Finally, section 2.6 defines focus measures.

2.1 Hyperspectral images and optics

Hyperspectral cameras have hundreds of spectral bands, typically in the 400-2000 nm range. This information makes it possible to distinguish between materials that may appear similar in RGB-images. This is because different materials have different absorption characteristics, quantified by the materials absorption coefficient. The absorption coefficient (typically denoted μ_a) of a material is defined as the probability per unit length that a photon is absorbed by the material. It has unit m^{-1} and depends on the wavelength of the light[12].

When capturing hyperspectral images of an object, light travels through the object, and is absorbed in varying amounts at different wavelengths according to the absoption coefficient. Light exits the object and is captured by the camera. The captured light can then be compared with absorption coefficients of known materials to identify concentrations of different materials and composition within the object. These process is relatively simple when absorption is the only light-matter interaction present. This is not the case in turbid media, where light scattering is the dominant light-matter interaction.

Scattering is the interaction between photons and matter where photons change trajectory in a random manner. The scattering characteristics of a material is quantified by the scattering coefficient (typically denoted μ_s), defined as the probability per unit length that a photon is scattered by the material, and the scattering anisotropy (typically denoted g) defined as the average cosine of the polar scattering angle[12].

When light travels through turbid medium, photons that enter the media together scatter independently of each other, and are spread out throughout the medium, as shown in the middle of figure 1. This behaviour is described by the point-spread function (PSF), or the impulse response of the system. In the case of a medium with constant thickness, the PSF is a function $P(\boldsymbol{\rho}) : \mathbb{R}^2 \to \mathbb{R}$ where the function value describes the intensity of light at a point ρ on the exit surface of the medium that results from an impulse of light entering the medium at the origin. The PSF naturally depends on the optical and geometrical properties of the medium. Since the optical properties of a medium typically are wavelength

dependent, the same is the case for the PSF. This is illustrated in the right plot of figure 1.

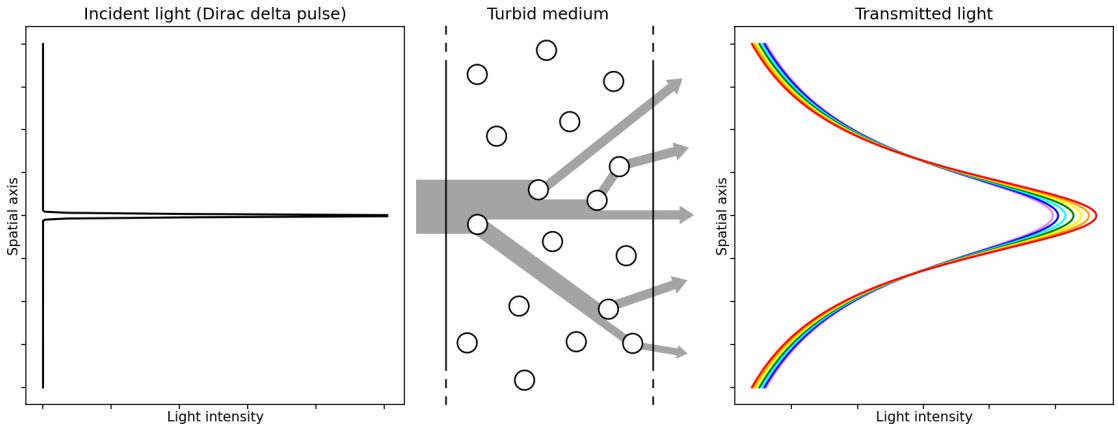


Figure 1: Scattering of a Dirac delta pulse illustrated in a one-dimensional space.

2.2 Boltzmann's transport equation and diffusion theory

To accurately quantify optical properties of a medium under observation, we need a mathematical framework for describing light propagation. There are many to choose from. Maxwell's equations fully describe the propagation of light as waves. However, if the geometries in question are complex, solving these are infeasible[12]. Another option is to simulate each individual photon using a Monte Carlo simulation. This method as been used successfully in many applications[20], but is very computationally demanding. A third approach is to consider the transfer of energy, described by Boltzmann's transport equation, shown in (1). Table 1 lists the symbol definitions for the equation as given by [12].

$$\frac{1}{c}\frac{\partial}{\partial t}L(\boldsymbol{r},\boldsymbol{s},t) = -\nabla\cdot[\boldsymbol{s}L(\boldsymbol{r},\boldsymbol{s},t)] - \mu_t L(\boldsymbol{r},\boldsymbol{s},t) + \mu_s \int_{4\pi} L(\boldsymbol{r},\boldsymbol{s},t)p(\boldsymbol{s},\boldsymbol{s}')d\Omega' + Q(\boldsymbol{r},\boldsymbol{s},t) \tag{1}$$

Table 1: Symbol definitions for equation (1) as given by [12]

Quantity	**Explanation**	**Unit**
c	The speed of light in the medium	m/s
μ_a	Probability per unit length traveled that a photon is absorbed by the material	m^{-1}
μ_s	Probability per unit length traveled that a photon is scattered by the material	m^{-1}
g	Average cosine of the polar scattering angle	Unitless
μ_s'	$\mu_s' = \mu_s(1-g)$	m^{-1}
μ_t	$\mu_t = \mu_a + \mu_s'$	m^{-1}
$I(\boldsymbol{r}, \boldsymbol{s}, t)$	The radiant intensity I at a point $\boldsymbol{r}$ in the direction $\boldsymbol{s}$ is the radiant power dP leaving the source in a an element of solid angle $d\Omega$ divided by $d\Omega$: $I = \frac{dP}{d\Omega}$	W/sr
$L(\boldsymbol{r}, \boldsymbol{s}, t)$	The radiance L at a point $\boldsymbol{r}$ on a surface in the direction $\boldsymbol{s}$ is the radiant intensity dI of an element dA of the surface, divided by the area of the orthogonal projection of dA on the plane orthogonal to $\boldsymbol{s}$: $L = \frac{dI}{dA\cos\theta}$	W/(sr$\cdot m^2$)
$p(\boldsymbol{s}, \boldsymbol{s}')$	The phase function defining the probability density that a photon is scattered from the direction $\boldsymbol{s}'$ to the direction $\boldsymbol{s}$	Unitless
$Q(\boldsymbol{r}, \boldsymbol{s}, t)$	The source function defining the radiant intensity of the point $\boldsymbol{r}$ in the direction $\boldsymbol{s}$	W/sr

Boltzmann's transport equation is derived based on the conservation of energy. The left side represents the change over time in the radiance L within a differential volume. This change must equal the sum of the gain-, and loss mechanisms on the right hand side. A physical interpretation of each of the gain-, and loss mechanisms are given in table 2.

Table 2: Physical interpretations of the rhs. terms of (1). The equation states that the change in the radiance L within a differential volume must equal the sum of these gain-, and loss mechanisms.

Term	**Explanation**	**Gain/Loss**
$-\nabla \cdot [\boldsymbol{s}L(\boldsymbol{r}, \boldsymbol{s}, t)]$	The power lost due to the energy gradient. Energy flows from areas with high potential to areas with low potential	Both
$-\mu_t L(\boldsymbol{r}, \boldsymbol{s}, t)$	The power lost from the volume due to absorption and scattering	Loss
$\mu_s \int_{4\pi} L(\boldsymbol{r}, \boldsymbol{s}, t) p(\boldsymbol{s}, \boldsymbol{s}') d\Omega'$	The gain in power due to scattering events into the volume	Gain
$Q(\boldsymbol{r}, \boldsymbol{s}, t)$	The source intensity of the volume	Gain

In the steady-state case, (1) reduces to (2)[12].

$$0 = -\nabla \cdot [\boldsymbol{s}L(\boldsymbol{r}, \boldsymbol{s})] - \mu_t L(\boldsymbol{r}, \boldsymbol{s}) + \mu_s \int_{4\pi} L(\boldsymbol{r}, \boldsymbol{s}) p(\boldsymbol{s}, \boldsymbol{s}') d\Omega' + Q(\boldsymbol{r}, \boldsymbol{s}) \tag{2}$$

When scatting is the dominant light-matter interaction, the radiance $L(r, s)$ can be approximated as being close to isotropic. This is because after sufficient scattering events, light inside the medium loose it's directionality. For this approximation to be accurate $\mu_s' >> \mu_a$ must be satisfied. The approximation is called the diffusion approximation, and gives the following expression for the radiance[12].

$$L = \frac{1}{4\pi}\phi + \frac{3}{4\pi}\boldsymbol{j} \cdot \boldsymbol{s} \tag{3}$$

where ϕ is the fluence rate and j is the energy flux.

Inserting (3) into (2) and integration over all solid angles we get the equation of continuity[12]:

$$\nabla \boldsymbol{j} = -\mu_a \phi + q \tag{4}$$

where q is the radiant power of the source.

Together with Fick's law, shown in (5)[12], the equation of continuity gives the diffusion equation, shown in (6)[12]:

$$\boldsymbol{j} = -D\nabla\phi \tag{5}$$

where $D = 1/(3\mu_t)$ is the diffusion constant.

$$\nabla^2\phi - \frac{\phi}{\delta^2} = -\frac{q}{D} \tag{6}$$

where $\delta^2 = 1/(3\mu_a\mu_t)$.

2.3 Signals and signal notation

A signal is a function $f : \mathbb{A} \to \mathbb{B}$. $\mathbb{A}$ can be single- or multi-dimensional, and the signal is classified accordingly. If $\mathbb{A} = \mathbb{N}^n$, the signal is called discrete. If $\mathbb{A} = \mathbb{R}^n$, the signal is called continuous. As the focus of this thesis is images, the focus will be on two-dimensional discrete signals $f : \mathbb{N}^2 \to \mathbb{R}$. Signal values will be denoted by $f[m, n]$.

In practice, image signals always has a finite domain. Unless otherwise specified this will be $\mathbb{A} = \{(x, y) \in \mathbb{N}^2 \mid 0 \leq x < M, 0 \leq y < N\}$. However, for mathematical convenience, this thesis will view images as infinite signals, with domain $\mathbb{N}^2$, where $f[m, n] = 0$ when $(m, n) \notin \mathbb{A}$. Similar notation and convenience will be used for one-dimensional signals.

2.4 Denoising

Denoising is the process of removing noise from a signal. This is useful in a wide variety of applications, including all forms of imaging. No capture system is perfect, so the captured signal is always affected by noise to some degree. In the case of additive noise, the observed signal f is modeled as a sum of the desired signal g and a noise signal n:

$$f = g + n \tag{7}$$

Denoising is the process of recovering g from f without absolute knowledge of n. To do this, one must assume a model for n. For many applications, the noise is assumed to be normally distributed with zero mean and some variance σ_n:

$$n \sim \frac{1}{\sigma_n\sqrt{2\pi}} e^{-\frac{1}{2}x^2/\sigma_n^2} \tag{8}$$

There are many methods for solving this type of problem. Two of them will be presented below.

2.4.1 Total variation (TV) denoising

Total variation (TV) denoising is based on the assumption that noisy signals have high *total variation*, defined as

$$V(g) = \sum_{n \in \mathbb{N}} |g[n+1] - g[n]| \quad \text{(1D TV)} \tag{9}$$

$$V(g) = \sum_{m \in \mathbb{N}} \sum_{n \in \mathbb{N}} \sqrt{(g[m+1,n] - g[m,n])^2 + (g[m,n+1] - g[m,n])^2} \quad \text{(2D TV)} \tag{10}$$

TV denoising reconstructs g from f by minimizing the total variation in g while minimizing the square error between g and f, $E(f,g)$:

$$\hat{g} = \arg\min_{g} \{E(f,g) + \lambda V(g)\} \tag{11}$$

In the one-, and two-dimensional cases $E(f,g)$ would be

$$E(f,g) = \frac{1}{N} \sum_{n \in \mathbb{N}} (g[n] - f[n])^2 \quad \text{(1D error)} \tag{12}$$

$$E(f,g) = \frac{1}{MN} \sum_{m \in \mathbb{N}} \sum_{n \in \mathbb{N}} (g[m,n] - f[m,n])^2 \quad \text{(2D error)} \tag{13}$$

The regularization term λ controls how "agressive" the denoising is. $\lambda = 0$ gives no denoising. As $\lambda \to \infty$ the reconstructed signal $\hat{g}$ becomes progressively smoother, and less like the original signal f. Figure 2 and 3 show examples of TV denoising on a one-dimensional and two-dimensional signal, respectively.

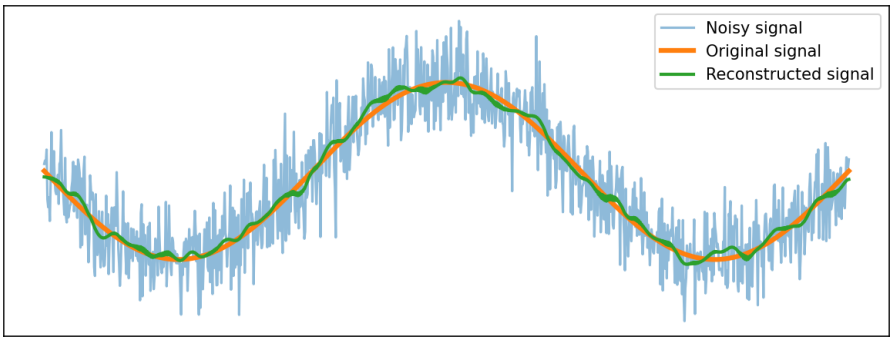
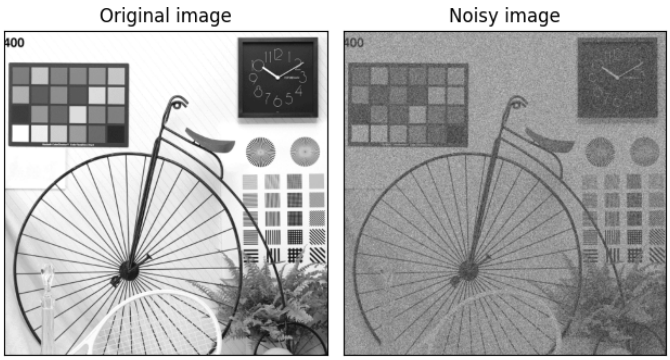
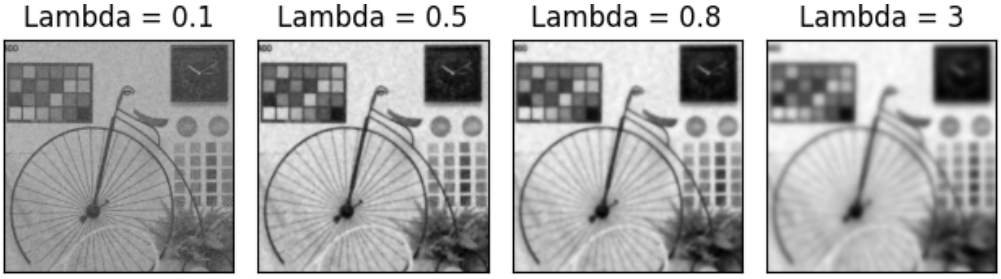


Figure 2: Example of TV denoising on a one dimensional signal.



(a) Left: Noisless test image, right: test image with additive gaussian noise.



(b) Results after TV denoising with varying values for λ.

Figure 3: Example of TV denoising on a two dimensional signal. A test image (top left) is corrupted by gaussian noise and then restored using TV denoising using varying values for λ.

2.4.2 Denoising by block-matching and 3D filtering (BM3D)

The BM3D algorithm is a recent and very popular algorithm for image denoising. It was first proposed by Dabov et. al.[2] in 2007. The algorithm can be broken down into the following steps:

1. **Block-matching:**
   Find blocks of the image that are similar, and group them into 3D-blocks.

2. **Collaborative filtering of the blocks:**
   Apply a linear 3D transform to the blocks. Then, remove information from the transformed coefficients, for example using thresholding.

3. **Reconstruction:**
   Reconstruct the image blocks by applying the inverse of the transform in step 2 and reassemble the image from the reconstructed blocks.

The algorithm is based on the assumption that most natural images have a high degree of spatial redundancy, meaning that different parts of the image can be represented as a combination of the same basic elements. This assumption implies that when the transform domain is shrunk (step 2), the removed information is mostly noise, while the true image is kept. Details about each step is presented below.

Block-matching

The block-matching approach used in the BM3D algorithm is the same as the one used in MPEG 1, 2, and 4, and H.26x. Each group of blocks is created in respect to a reference block. Every block that is similar enough to this reference block is included in the group. Similarity is measured using a distance measure like an l^p norm. This approach produces blocks that are not necessarily disjoint, meaning that the same block can be part of multiple groups. Figure 4 shows an example of block matching in practice.

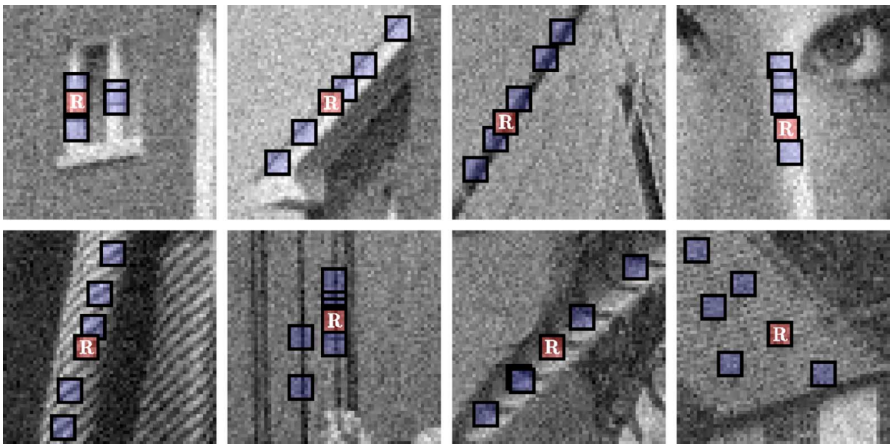


Figure 4: Retrieved from [2]. Illustration of grouping blocks from noisy natural images corrupted by white Gaussian noise with standard deviation 15 and zero mean. Each fragment shows a reference block marked with"R"and a few of the blocks matched to it.

Collaborative filtering of the blocks

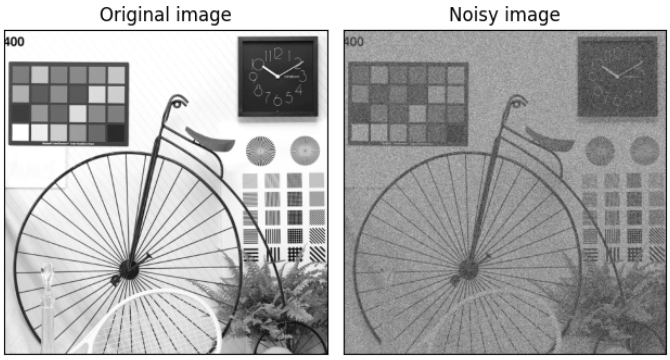
Because the blocks in each group are similar, transforming the group by some linear transform means that we can represent each block by a small number of coefficients in the transform domain. Because of our assumption that natural images have a high degree of spatial redundancy, the largest coefficients should contain the true image information, while the smaller coefficients should mostly contain information about the noise. Therefore, shrinking the transform domain

is an effective way of denoising the image. In this step, the σ parameter of the algorithm controls how aggressive the denoising is, by varying the amount the transform domain is shrunk.

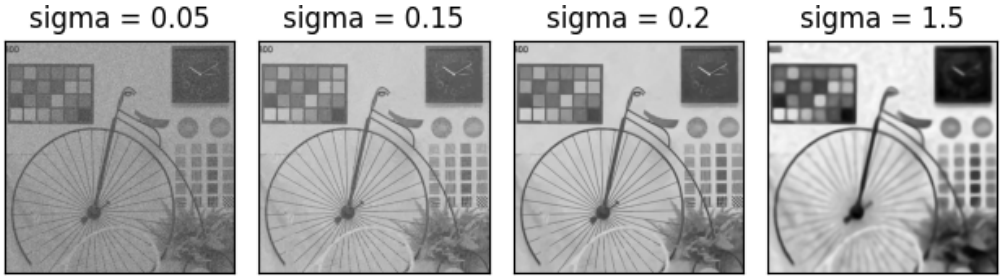
Reconstruction

After shrinking the transform domain, the blocks are reconstructed by applying the inverse transform of the one used in the previous step. The blocks are the placed back into their original location. Because blocks may overlap, and that the groups are not necessarily disjoint, many pixels will have multiple estimates from different blocks. These estimates are aggregated to produce the best possible reconstruction

Figure 5 shows an example of the results from the BM3D algorithm.



(a) Left: Noisless test image, right: test image with additive gaussian noise.



(b) Results after BM3D denoising with varying values for σ.

Figure 5: Example of BM3D denoising. A test image (top left) is corrupted by gaussian noise and then restored using BM3D using varying values for σ.

2.5 Convolution and deconvolution

Convolution is a mathematical operation on two functions or signals. The operation comes in both continuous and discrete verities, as well as single- and

multidimensional signals. The convolution of g with f is denoted by $g * h$. Since the subject of this thesis is image processing, $*$ will always denote discrete 2D-convolution, defined by

$$(g * h)[i, j] = \sum_{m \in \mathbb{N}} \sum_{n \in \mathbb{N}} g[i - m, j - n] \cdot h[m, n] \tag{14}$$

Given a signal $f = g * h$, deconvolution is the process of recovering g from f and h. The following sections will describe two deconvolution algorithms.

2.5.1 Weiner deconvolution

Both convolution and deconvolution is commonly done in the Fourier domain. Let F, G and H be the Fourier transforms of f, g and h, respectively. Then $f = g*h$ can be expressed in Fourier domain as $F = G \cdot H$ where $\cdot$ denotes regular multiplication. If we assume $H \neq 0$ on the entire domain, we can reconstruct G from F and H by

$$G = \frac{F}{H} \tag{15}$$

There are two problems with using this approach. The first is that the assumption $H \neq 0$ may not hold. The second problem is that in most real-world applications, we can't observe f. Instead we observe a noisy f, usually modeled by $f+n$ where n is additive noise, with Fourier transform N. Substituting $F+N$ for F in (15), we see that our recovered signal $G = F/H + N/H$ becomes very noisy when H is low.

Multiple methods have been developed to combat these problems, with one being Weiner deconvolution. The Weiner deconvolution of f with h finds the estimate, $\hat{g}$ of g that minimized the expected square error between $\hat{g}$ and g assuming that the noise is Gaussian. This is done based on knowledge of the signal-to-noise ratio (SNR), defined by $SNR = \frac{S_{ff}}{S_{nn}} = \frac{\mathbb{E}|F|^2}{\mathbb{E}|N|^2}$.

The Weiner deconvolution of f with h is expressed in Fourier domain by (16)[18].

$$\hat{G} = \frac{F}{H}\left(1 + \frac{1}{|H|^2 SNR}\right)^{-1} \tag{16}$$

Even when the SNR is unknown, as is the case in most applications, the method can produce good results by using a constant estimate for the SNR.

2.5.2 Richardson–Lucy deconvolution

The Richardson-Lucy deconvolution algorithm is an iterative algorithm designed to reconstruct blurry emission tomography (ET) images[10]. The observed signal

f is modeled as $f = \widetilde{g} * h$ where $\widetilde{g}$ follows a Poisson distribution with mean g. This model is chosen to fit the process of capturing of ET images, but the algorithm as been shown to be useful in more general applications[4].

Given an initial estimate $\hat{g}_0$ of g, the algorithm produces a series of estimates $\hat{g}_i$ where each estimate is calculated from the previous estimate by (17).

$$\hat{g}_{i+1} = \hat{g}_i \cdot \left[\bar{h} * \frac{f}{h * \hat{g}_i} \right] \tag{17}$$

where multiplication and division is done element-wise and $\bar{h}$ is the flipped version of h. When h is symmetric $\bar{h} = h$.

The algorithm has two important properties. Firstly, $p(f|\hat{g}_{i+1}) \geq p(f|\hat{g}_i)$, meaning that each estimate is better than or equal to the previous estimate. Secondly, $\hat{g}_i$ approaches the maximum likelihood estimate of g as $i \to \infty$. The derivation for this is long will not be given here, but can be found in [10]. An intuitive explanation of why (17) produces these properties will be given below.

When $\hat{g}_i = g$, $h * \hat{g}_i = f$ and the fraction in (17) is 1. Convolving this with $\bar{h}$, we get 1, and $\hat{g}_{i+1} = \hat{g}_i$, meaning that $\hat{g}_i = g$ is a steady state solution to (17). When $h * \hat{g}_i \neq f$, the fraction is large in places where $h * \hat{g}_i$ is too low compared to f. In the case of symmetric h, the fraction is convolved with h, meaning that the factor inside the parenthesis of (17) is large at and around places where an increase in $\hat{g}_i$ would make $h * \hat{g}_i$ closer to f, and thereby $\hat{g}_i$ closer to g. Therefore, applying (17) to an estimate $\hat{g}_i$ nudges it closer to the true g.

2.6 Focus measures

A focus measure is a function $\alpha : \mathbb{I} \to \mathbb{R}$ where $\mathbb{I}$ is the space of image signals. A focus measure quantifies the sharpness of an image, meaning that if $\alpha(f_a) < \alpha(f_b)$, then f_b is considered to be sharper (or equivalently, less blurry) than f_a. Focus measures are widely used for to do auto-focus in digital photography, and are also used in some of the methods explored in this thesis.

In an article by Mir et. al.[8] 35 focus measures were evaluated in context of the application of digital photography. In an article by Sun et. al.[14] 18 focus measures was evaluated in the context of computer microscopy. Two of these measures will be presented below.

The Brenner focus measure for an image f is defined by

$$\alpha_{Brenner}(f) = \sum_{m \in \mathbb{N}} \sum_{n \in \mathbb{N}} (f[m, n+2] - f[m, n])^2 \tag{18}$$

Similarly the vertical Brenner focus measure is defined by

$$\alpha_{Brenner,v}(f) = \sum_{m \in \mathbb{N}} \sum_{n \in \mathbb{N}} (f[m+2, n] - f[m, n])^2 \tag{19}$$

The Brenner focus measures can be view as edge detectors. They detect presence of edges and sum the contributions from the entire image. The underlying assumption is that more and harder edges are an indicator of a more in-focus image. The Brenner focus measure performed well in the evaluation done by Mir et. al. and average in the evaluation by Sun et. al.

Image statistics can also be used as a focus measure. The normalized variance of an image f is defined as

$$\alpha_\sigma(f) = \frac{1}{NM\mu} \sum_{m \in \mathbb{N}} \sum_{n \in \mathbb{N}} (f[m,n] - \mu)^2 \tag{20}$$

with

$$\mu(f) = \sum_{m \in \mathbb{N}} \sum_{n \in \mathbb{N}} f[x,y] \tag{21}$$

The normalized variance focus measure preformed well in the evaluation by Sun et. al. and average in the evaluation by Hasim et. al.

3 Literature review

A literature review was performed to identify methods that can help in reaching this thesis' goal. The literature is rich with methods for deblurring. These methods vary in how much knowledge of the PSF is assumed.

Some methods, such as the ones in [10] and [13], assume that the PSF is known, and try to optimize the reconstruction with this knowledge. Other methods, such as the one in [7]. Don't assume knowledge of the PSF, but rather statistical knowledge of the characteristics of a reconstructed image. Machine learning has also been used to effectively deblur images. In [19], Zhao et. al. use blurry images and non-blurry images to train an machine learning model to deblur images without explicit knowledge of the PSF.

To obtain a manageable scope for this thesis, detailed literature review and implementation is only done for methods that assume blurring as a result of scattering through turbid medium. This is not, however, limited to methods that assume hyperspectral data. A table of the most relevant methods found by the literature review process is shown below, with details given in the following sections.

Author(s)	**Application**	**Working principle**	**Input data**
Shimizu et. al.	Functional, fluorescent imaging of rats	Deconvolution with parameterized PSF. Parameters fitted using focus-finding principles.	Single channel fluorescent image
Lee and Park	Authentication by finger vein recognition	Deconvolution with parameterized PSF. Parameters fitted using statistical knowledge of the desired image.	Single channel image captured at 870nm
Muria and Sato	Authentication by finger vein recognition	Reconstruction of deblurred image by exploiting that channels at different wavelengths are blurred in different and predictable amounts.	Three channel image
Jiang et. al.	Spiral CT imaging of the inner ear	Deconvolution with parameterized, gaussian PSF. Parameters fitted using ENR measure	CT images
Svaasand et. al., Bjorgan et. al.	Estimation of skin optical parameters	Inverse modeling and diffusion theory	NIR hyperspectral data cube

Table 3: A summary of methods from the literature that is chosen for in-depth exploration and implementation.

3.1 Methods using deconvolution with a parameterized PSF

Methods for deblurring by deconvolution with a parameterized PSF are common in the literature. These models assume that the captured image I', weather it be a CT image, fluorescence image or NIR image, can be modeled as a convolution between the desired, clear image I and a PSF with additive noise ϵ, as show in (22). The PSF is assumed to be a known function of unknown parameters. The way the parameters are determined vary between methods. When the parameters of the PSF are determined, the deblurred image can be found by deconvolving the captured image with the PSF.

$$I' = I * PSF + \epsilon \tag{22}$$

In [11], Shimizu et. al. develop a method for improving fluorescent images. Shimizu et. al. uses diffusion theory of photon transport to derive a closed form PSF for a homogeneous turbid medium of constant thickness. The PSF and is shown in (23) and depends on the absorption coefficient, the reduced scattering coefficient μ_s', the thickness d of the medium and the total power of a point source P_0.

$$PSF(\boldsymbol{\rho}; d, P_0) = \frac{3P_0}{4\pi^2}\left\{(\mu_s' + \mu_a) + \left[\delta + \frac{1}{l}\right]\frac{d}{l}\right\}\frac{e^{-\delta l}}{l} \tag{23}$$

With $l^2 = \boldsymbol{\rho}^2 + d^2$ and $\delta^2 = 3\mu_a(\mu_s' + \mu_a)$.

μ_a and μ_s' are assumed known. To determine d_0 and P_0, Shimizu et. al. use a focus-finding algorithm. By selecting a focus measure $\alpha(I)$, one can define a function β mapping parameter values to the focus value of the reconstructed image, shown in 24. We will call this a *focus curve*.

$$\beta(\boldsymbol{\theta}) = \alpha(I' *^{-1} PSF(\boldsymbol{\theta})) \tag{24}$$

where $*^{-1}$ denotes deconvolution, and $\boldsymbol{\theta}$ is the PSF parameters, which in this case are d_0 and P_0. Then, determining the parameters d and P_0 can be done by finding the global maximum of $\beta(d, P_0)$:

$$(\hat{d}, \hat{P_0}) = \underset{(d,P_0)}{\arg\max}\{\beta(d, P_0)\} \tag{25}$$

With these estimates, an estimate of the clear image can be constructed by

$$\hat{I} = I' *^{-1} PSF(\hat{d}, \hat{P_0}) \tag{26}$$

Reconstructing in this way optimizes the focus of the recovered image.

Fluorescent images are different from the transmission images researched in this thesis in that the captured light originated inside the tissue, rather than traveling through it. However, in both cases, there is an optical signal embedded in the tissue, that is distorted by scattering as it travels through the tissue. The physics governing this scattering is the same in the two cases. Therefore, it is reasonable to assume that the method developed by Shimizu et. al. also can improve the quality of hyperspectral images.

Lee and Park[6] showed that it was possible to improve the quality of transmission images of finger veins captured by a webcam using a similar method. They used gray-scale images of the middle part of each finger, excluding thumbs, captured at 850 nm. They use the same parameterized PSF, but a different method for fitting the parameters d and P_0. This process starts by creating a map of where in the image blood vessels are located. Then, the orthogonal profile of every point of every vessel is captured. If the image was unaffected by

blurring, these profiles would be sharp. The parameters of the PSF are found by determining the parameter values that produces the best match between the captured profiles and a convolution between an idealized, sharp vessel profile, and the PSF at these parameters. This method is dependent on the assumption that all vessels in the image is of similar thickness. Based on where on the finger the images where captured, Lee and Park assume that all vessels in the image are close to 1.5 mm thick.

Deconvolution with a parameterized PSF is also the main idea behind the method developed by Jiang et. al. in [4]. In the paper, the autors develop a method for deblurring CT images of the inner ear. They assume a gaussian PSF an fit the parameters by maximizing the edge-to-noise ratio, defined by 27, of the deconvolved image. This is in principle the same procedure as the one employed by Shimizu et. al. and Lee and Park, but with a different focus measure and PSF.

$$ENR = \frac{E}{N} = \frac{d_K(\hat{I}, PSF * \hat{I})}{d_K(I, PSF * \hat{I})} \tag{27}$$

where d_K is the Kullback distance defined by

$$d_K(u, v) = \sum u \log \frac{u}{v} - \sum (u - v) \tag{28}$$

3.2 Deblurring using tri-band Illumination by Muria and Sato

In [9], Miura and Sato develop a method for deblurring 3-band transmission images of finger veins. The method exploits the fact that the scattering properties of tissue is different at different wavelengths. Miura and Sato models the captured image I'_i on each band i by a clear image I_i convolved with a band-dependent PSF:

$$I'_i = I_i * PSF_i + n_i \tag{29}$$

The foundational hypothesis of this method goes as follows. Choose two wavelengths λ_i and λ_j such that the corresponding extinction coefficients of skin $\mu_{tr,i}$ and $\mu_{tr,j}$ satisfy $\mu_{tr,i} = 2\mu_{tr,j}$. Then $PSF_i \approx PSF_j * PSF_j$. In other words, if the extinction coefficient at channel i is twice that of channel j, the captured image on channel i is "twice as blurry" as that on channel j.

Using this hypothesis, Miura and Sato show that a clear image $\hat{I}_j$ can be reconstructed from the two blurry images by

$$\hat{I}_j = I'_j + (I'_j - I'_i) * I'_j *^{-1} I'_i \tag{30}$$

Where $*^{-1}$ denotes deconvolution.

In the capturing process, Miura and Sato dynamically adjusts the illumination brightness, such that the three bands are of equal brightness. To preform the deconvolution, a Weiner filter is used. Smoothing of the second term in (30) is done to increase noise robustness. The entire reconstruction is done in patches of size 1x1 mm.

3.3 Scattering correction using diffusion theory

A popular way of correcting for the scattering of light is to use diffusion theory. In [15] and [1], Svaasand et. al. and Bjorgan et. al. use diffusion theory to obtain an expression for the reflectance of a layered medium in terms of it's optical parameters. This expression is used to estimate the optical parameters from the observed reflectance using inverse modeling. A good estimation of the absorption coefficient μ_a will naturally separate the absorption and scattering characteristics of a material, resulting in a scattering correction method.

This method of estimating μ_a was used in the project preceding this thesis. The following section is a derivation the reflectance model for a single layer medium obtained by diffusion theory. The derivation is from the preceding project's report[16] and is based primarely on [15]. A complete derivation, including all intermediate steps can be found in the appendix of [16].

3.3.1 Reflectance model

The starting point for obtaining a model for the reflectance of a sample will be the time-independent diffusion equation for photon transport, shown in (31) where ϕ is the fluence rate, q is the source function, $D = 1/(3\mu_{tr})$ is the diffusion constant and $\delta = 1/\sqrt{3\mu_{tr}\mu_a}$ is the optical penetration depth[15].

$$\nabla^2\phi - \frac{\phi}{\delta^2} = -\frac{q}{D} \tag{31}$$

Assuming homogeneity along x and y, the ∇^2 reduces to $\frac{\partial^2}{\partial z^2}$. Taking the laplace transform of (31) and solving for $\phi(z)(s)$ we get

$$\phi(z)(s) = \frac{sC_1}{s^2 - \frac{1}{\delta^2}} + \frac{C_2}{s^2 - \frac{1}{\delta^2}} - \frac{q}{D\left(s^2 - \frac{1}{\delta^2}\right)} \tag{32}$$

Assuming isotropic source functions, the photon source q in a homogeneous infinite medium can be described by (33)[15], with the laplace transform shown in (34).

$$q(z) = \mu_s' e^{\mu_{tr} z} \tag{33}$$

$$q(z) = \frac{\mu_s'}{s + \mu_{tr}} \tag{34}$$

Inserting (34) into (32) and solving for ϕ, yields the fluence rate for $z > 0$, ϕ^+ expressed in terms of unknown parameters C_5 and C_6:

$$\phi^+(z) = C_5 e^{\frac{z}{\delta}} + C_6 e^{-\frac{z}{\delta}} + \frac{\delta^2 \mu_s'}{D(1 - \mu_{tr}^2 \delta^2)} e^{-\mu_{tr} z}$$

The total energy in the system must be finite. Therefore, $\int_0^\infty \phi^+(z) dz$ must be bounded. This implies that $\lim_{z \to \infty} \phi^+(z) = 0$. Using this boundary condition, we find that $C_5 = 0$, reducing (35) to:

$$\phi^+(z) = C_6 e^{-\frac{z}{\delta}} + \frac{\delta^2 \mu_s'}{D(1 - \mu_{tr}^2 \delta^2)} e^{-\mu_{tr} z} \tag{35}$$

To determine C_6, the boundary condition for the skin-air interface shown in (36)[15] is used.

$$j(z = 0) = A \phi^+(z = 0) \tag{36}$$

where the coefficent A is defined in terms of the effective reflection coefficient R_{eff} as shown in (37)[15][3]. The complete definition of R_{eff} can be found in the appendix of [16].

$$A = \frac{1 - R_{eff}}{2(1 + R_{eff})} \tag{37}$$

Solving (36) yields

$$C_6 = \frac{A - D\mu_{tr}}{\frac{D}{\delta} - A} \frac{\delta^2 \mu_s'}{D(1 - \mu_{tr}^2 \delta^2)} \tag{38}$$

Inserting (38) into (35) yields an expression for the fluence rate in terms of assumed known variables:

$$\phi^+(z) = \frac{A - D\mu_{tr}}{\frac{D}{\delta} - A} \frac{\delta^2 \mu_s'}{D(1 - \mu_{tr}^2 \delta^2)} e^{-\frac{x}{\delta}} + \frac{\delta^2 \mu_s'}{D(1 - \mu_{tr}^2 \delta^2)} e^{-\mu_{tr} z} \tag{39}$$

Using (39), the reflectance γ is calculated using (40)[15]

$$\gamma = j(z = 0) \tag{40}$$

$$\gamma = \frac{A\delta^2\mu_s'}{(D - \delta A)(1 + \mu_{tr}\delta)} \tag{41}$$

Rewriting D, δ and μ_{tr} in terms of μ_a and μ_s' yields a reflectance model depending only on the absorption coefficient μ_a the reduced scattering coefficient μ_s' and A.

$$\gamma = \frac{A\mu_s'}{(1 + A)\mu_a + A\mu_s' + \left(\frac{1}{3} + A\right)\sqrt{3\mu_s'\mu_a + 3\mu_a^2}} \tag{42}$$

4 Data and methods

Based on the findings from the literature review, tree methods were developed and applied on a dataset of hyperspectral images. A presentation of the dataset is given in section 4.1. The methods are described in section 4.2 - 4.4.

4.1 Agar dataset

The dataset consists of hyperspectral images of printed shapes covered by a Petri dish containing varying amounts of turbid media. The images were captured with a HySpex VNIR-1600 and contain spectral information from 160 wavelength bands ranging from 411 nm to 989 nm. Gray-scale versions of the images used in this thesis are shown in figures 6 - 8. Figure 6 shows the ground truth image. A perfect deblurring algorithm would reconstruct the structures in this image from the blurry images. The dataset was created by Remi Johansen[5].

Figure 6: Hyperspectral image of all printed shapes used in the dataset covered by an empty Petri dish. This is the ground truth.

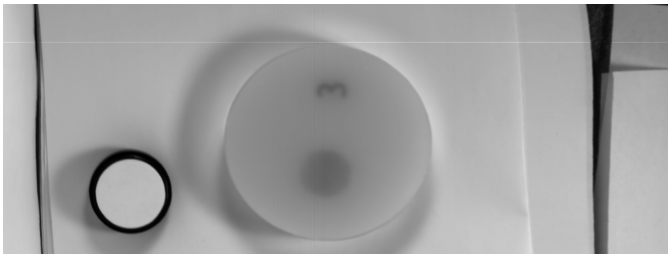


Figure 7: Hyperspectral image of the number three covered by a Petri dish with 3 mm turbid medium.

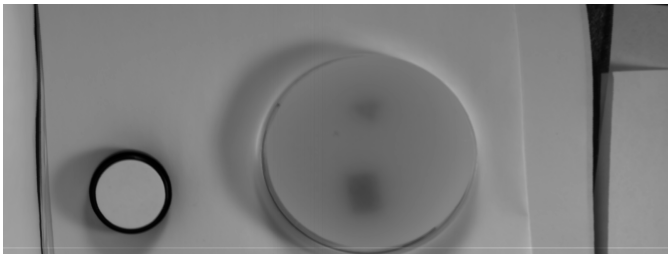


Figure 8: Hyperspectral image of the number four covered by a Petri dish with 5 mm turbid medium.

The medium is agar gel mixed with intralipid. Agar gel is in practice non-absorbing and non-scattering, while intralipid is practically non-absorbing and highly scattering. The resulting medium is therefore practically non-absorbing with the scattering properties of intralipid. Staveren et. al.[17] measured the scattering properties of 10% intralipid, and fitted the model shown in 43.

$$\begin{aligned} \mu_s &= 0.016\lambda^{-2.4} \\ g &= 1.1 - 0.58\lambda \end{aligned} \tag{43}$$

where λ is in micrometers and μ_s is in $mL^{-1}Lmm^{-1}$

The gels in the dataset have a 1.75% intralipid concentration. Estimating the scattering properties of this concentration by linear approximation gives the following expression for μ_s'

$$\mu_s(\lambda) = \frac{1.75}{10} 0.016\lambda^{-2.4}(1 - (1.1 - 0.58\lambda)) \tag{44}$$

All images in the dataset contain a Spectralon reflectance standard used for reflectance calibration. This is a material reflects over 99% of light in the 400 nm - 100 nm range[5]. Computing the average spectrum over the area of the reflectance standard gives a good estimate of the spectrum of the light source used. Each image was divided by it's light source spectrum to obtain reflectance values.

After calibration, the only intresting segment of the image is that covered by the Petri dish. This is extracted by cropping the image, and masking away the outside of the Petri dish as well as it's perimeter.

4.2 Method 1: Deconvolution with a parameterized PSF

This method is developed based on the work by Shimizu et. al [11] and Jiang et. al.[4]. The *focus curve* is a function of the parameter values $\boldsymbol{\theta}$ and is defined

as the value of the chosen focus measure on the deconvolution between the input image I' and the PSF with parameter values $\boldsymbol{\theta}$:

$$\beta(\boldsymbol{\theta}) = \alpha(I' *^{-1} PSF(\boldsymbol{\theta}))$$

The reconstructed image is the result of the deconvolution for the θ where the focus curve has it's maximum. Throughout this thesis the maximum will be found by sampling the focus curve at a sufficiently large interval with sufficient resolution, and selecting the largest value. For applications where shorter runtime is required, the gradient ascent algorithm can be used to find the maximum much more efficiently.

There are many ways to implement this general algorithm. There are multiple choices for the PSF, the deconvolution algorithm, the focus measure, and eventual preprocessing steps. For each of these four components, the options will be considered. The most promising options will be included in an evaluation where all combinations for each of the four components are tested.

For the PSF, the expression derived by Shimizu et. al is used. This is because from a physics perspective, their work is similar to this thesis. Additionally, the PSF used by Shimizu et. al was used successfully for finger vein deblurring by Lee and Park[6] on data similar to the data used in this thesis. One modification of the PSF is done to reduce complexity of the method. For the purposes of this thesis, there is no use in considering P_0 as a variable. P_0 appears as a multiplicative factor in the PSF. Therefore, variations in P_0 only result scaling of the image values resulting from the deconvolution. To remove the variable, the PSF is normalized with respect to P_0. This leaves the depth d as the only PSF parameter left to be fitted.

For the deconvolution algorithm, two options will be evaluated: Richardson-Lucy deconvolution, as used by Shimizu et.al., and Weiner deconvolution.

For the focus measure, three options will be evaluated: ENR as defined by Jiang et. al, the Brenner focus measure defined in (18), and the measure of normalized variance defined in (20). These three are chosen because they are all different in their approach to measuring focus. One could include all 35 focus measures described in [8], or all 18 described in [14]. However, since most of these are similar in nature to one of the three mentioned above, it is assumed that these would have similar performance.

As seen in section 2.5, deconvolution algorithms are sensitive to noise. Therefore, a denoising step prior to the deconvolution is included in the evaluation. BM3D is chosen as the denoising algorithm, with parameter value with $\sigma = 0.02$. BM3D is chosen because it has state of the art performance and is easy to implement in Python.

With one option for the PSF, two options for the deconvolution algorithm, three options for the focus measure, and two options for preprossesing, there are

12 possible configurations for the algorithm. The results of each configuration will be shown in section 5.1.

4.3 Method 2: Double blur reconstruction

This method is a reproduction of the method from Miura and Sato[9], with small modifications to accommodate the change in camera technology, and experimental setup.

Miura and Sato control the wavelength and intensity of the illumination elements in the experimental setup to capture a three band image at wavelengths λ_0, λ_1, λ_2, where all bands have equal average pixel brightness. This capturing process can be emulated by simply selecting the desired bands from the hyperspectral data used in this thesis. To achieve equal average pixel brightness, the bands are scaled in software.

The wavelengths λ_0 and λ_2 are chosen such that $\mu_s(\lambda_0) = 2\mu_s(\lambda_2)$, to satisfy the methods foundational hypothesis. Figure 9 shows a plot of (43) for the wavelength range of the hyperspectral data. The figure shows that the dataset includes a range of wavelengths that satisfy $\mu_s(\lambda_0) = 2\mu_s(\lambda_2)$. Visual inspection of the bands show that both the highest- and lowest wavelength bands lower signal to noise than the middle bands. Therefor 448 nm and 892 nm are chosen as λ_0 and λ_2, respectively. Corresponding bands I_0 and I_2 are extracted from the hyperspectral image and normalized such that all pixel values were in the range [0, 1]. The band with the highest average pixel brightness is scale down by a constant factor to match that of the other band. Then, BM3D denoising is used with $\sigma = 0.02$. This is because the algorithm uses deconvolution, which is sensitive to noise. After denoising, I_0 and I_2 are used to create a reconstructed image by applying (30). The deconvolution in (30) is implemented using Weiner deconvolution, as described in section 2.5.1. This is the same deconvolution algorithm as used by Miura and Sato.

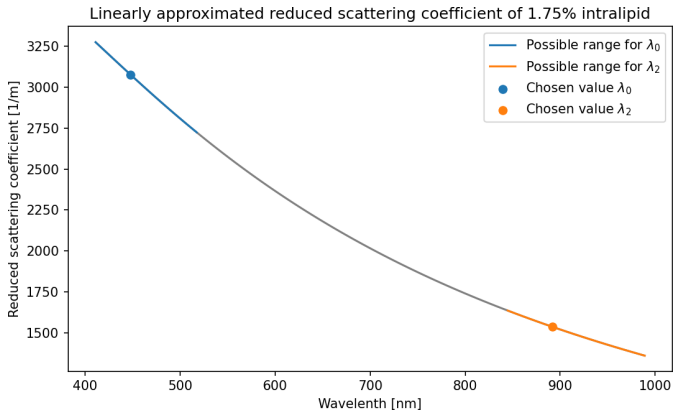


Figure 9: Plot of (43)

4.4 Method 3: Deblurring by absorption estimation

Method 3 is based on diffusion theory, and borrows many concepts from [1]. This framework has traditionally been used to remove the effects of scattering in the spectral domain, but this thesis will explore weather it can also be used to correct for spatial blurring.

Method 3 deblurres the hyperspectral image by looking at each pixel individually, and determining to what degree this pixel is part of the imaged structure. First, the absorption spectrum in each pixel is estimated from it's reflectance spectrum. This is done using the reflectance model derived in section 3.3.1. Details are given in section 4.4.1. Then, the correlation between the absorption spectrum and the structure's true spectrum is measured. This correlation is set as the pixel value in the reconstructed image. The structure's true spectrum is extracted from the ground truth image. Details are given in section 4.4.2.

4.4.1 Estimation of μ_a

Assuming knowledge of μ_s' and A, the reflectance model derived in section 3.3.1 gives the reflectance γ as a function of μ_a: $\gamma = f(\mu_a)$. Every sample in the calibrated hyperspectral data is a measurement of the reflectance $\gamma_{measured}$ at a certain point in space for a certain wavelength. An estimate for the μ_a for this point in space and wavelength can be obtained by finding the $\hat{\mu}_a$ that satisfies $f(\hat{\mu}_a) = \gamma_{measured}$. This can be done by using Newton's method to find zeros of the function $f(\hat{\mu}_a) - \gamma_{measured}$. This approach is used by Bjorgan et. al.[1] to fit multiple skin optical parameters. In the method used here, μ_a is the only parameter being estimated. Therefore, estimation can be done by numerically approximating f^{-1} and letting $\hat{\mu}_a = f^{-1}(\gamma_{measured})$.

f^{-1} is approximated numerically by sampling f and interpolating. First a list of $N = 1000$ evenly spaced values for μ_a is created: $[\mu_{a,1}, \mu_{a,2}, ..., \mu_{a,N}]$, with $\mu_{a,1} = 1m^{-1}$ and $\mu_{a,N} = 1000m^{-1}$. These values are all fed into f to create a list of 1000 corresponding reflectance values $\Gamma = [\gamma_1, \gamma_2, ..., \gamma_N]$. Then $f^{-1}(\gamma)$ is is approximated by linear interpolation:

$$f^{-1}(\gamma) = \mu_{a,i} + (\gamma - \gamma_i)\frac{\mu_{a,i+1} - \mu_{a,i}}{\gamma_{i+1} - \gamma_i} \tag{45}$$

where γ_i is the largest value in Γ such that $\gamma_i < \gamma$.

For μ_s', the approximation given in (43) is used. Since this is wavelength dependent, f^{-1} must be calculated for each band separately. For A, the value of 0.17 is used. This is the common value used for skin measurements[15][1], which is what the dataset used in this thesis is emulating. The resulting f^{-1} is plotted in figure 10.

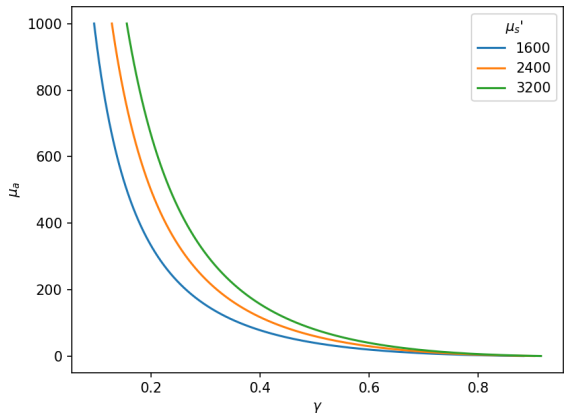


Figure 10: f^{-1} approximated by sampling of f and interpolating

4.4.2 Correlation with the true spectrum

To compute the structure's true spectrum $\mu_{a,true}[\lambda]$, the ground truth image was masked, setting all pixels not part of the number two, three or four to a value of zero. Then the average spectrum of the remaining pixels was computed.

After the absorption $\hat{\mu}_a[i, j, \lambda]$ is estimated for every pixel $[i, j]$ at every band λ, the corresponding pixel $I_{reconstructed}[i, j]$ in the reconstructed image is set as the correlation between the estimated absorption spectrum and the true spectrum:

$$I_{reconstructed}[i, j] = \sum_{\lambda} \hat{\mu}_a[i, j, \lambda] \cdot \mu_{a,true}[\lambda] \tag{46}$$

5 Results

The three methods described in section 4 were applied on the dataset. Sections 5.1-5.3 presents the results of the three methods.

5.1 Method 1

The 12 different configurations of the algorithm was applied on the two of the images in the agar dataset with 3 and 5 mm agar covering the structures. The focus curve for each configuration is computed for each image. The focus curves for the 3 and 5 mm case are shown in figure 11 and 12, respectively. Reconstructed images for different PSF parameter values for the 3 and 5 mm case are shown in figure 13 and 14, respectively.

Visual inspection of the reconstructed images at the bounds of the parameter intervals for the focus curves in figure 13 and 14 shows that the optimal depth is somewhere inside the interval. In the 3 mm case, only one of the 12 configurations have a maximum inside this interval. This is the configuration using Richardson-Lucy deconvolution with the ENR focus measure and BM3D denoising, similar to the algorithm used by Jiang et. al. The maximum of the focus curve is at $d = 3.11$. The algorithm output is the image corresponding to this depth. In the 5 mm case, none of the 12 configurations have a maximum inside this interval. The algorithm therefore fails to produce an output. When comparing the three methods, the image reconstructed with $d = 5$mm will be used as a stand-in for a real output from the algorithm.

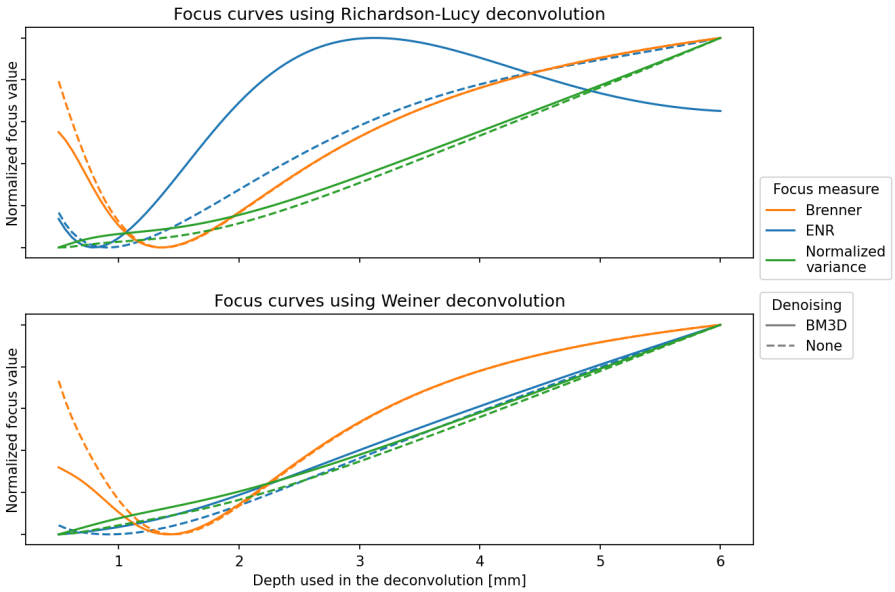


Figure 11: Focus curves for each of the 12 configurations of the algorithm when applied on an image of a structure covered by 3 mm turbid medium.

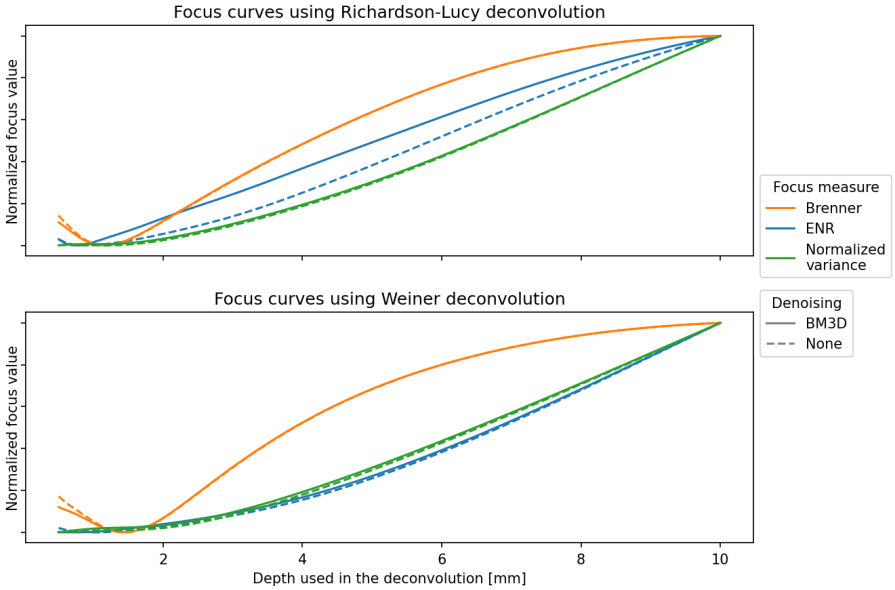


Figure 12: Focus curves for each of the 12 configurations of the algorithm when applied on an image of a structure covered by 5 mm turbid medium.

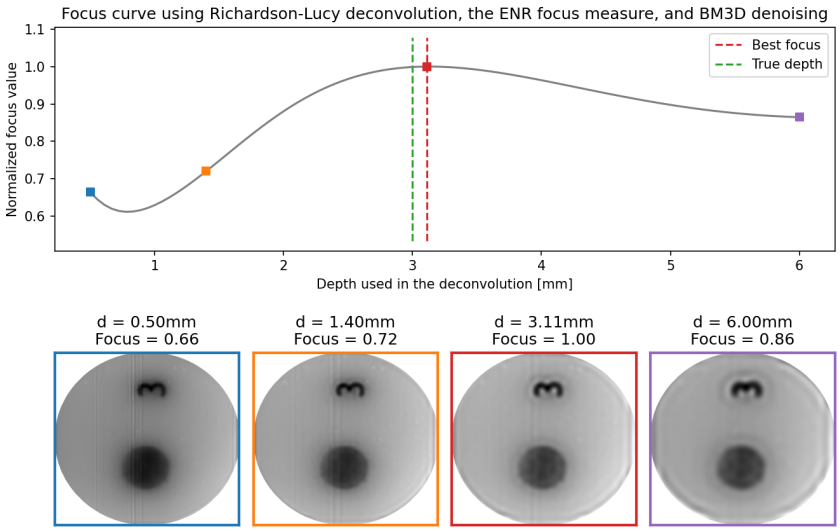


Figure 13: Top: Focus curve of the most promising configuration of the algorithm when applied on an image of a structure covered by 3 mm turbid medium. Bottom: Reconstructed images at different depths. The image with the highest focus is framed in red and is the algorithm output.

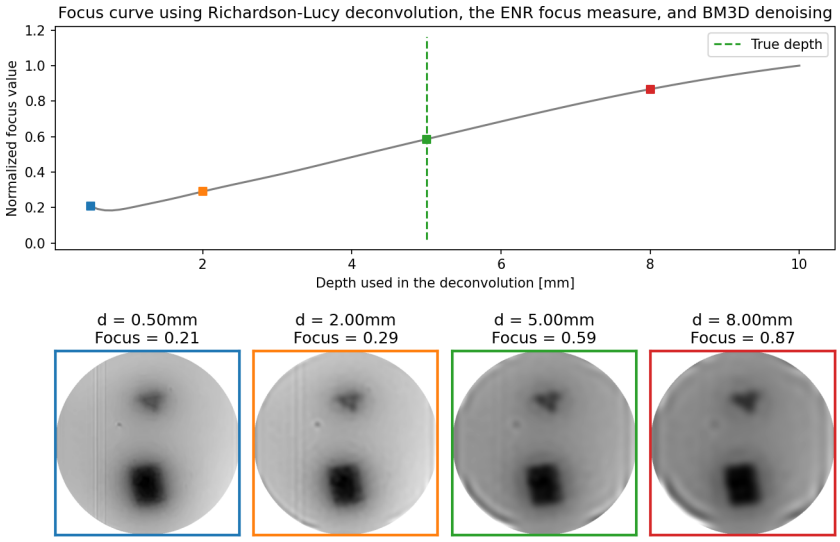


Figure 14: Top: Focus curve using the configuration that was most promising when applied on an image of a structure covered by 3 mm turbid medium, here applied on an image of a structure covered by 5 mm of turbid medium. Bottom: Reconstructed images at different depths. The focus curve has no maximum inside the interval. Therefore, the algorithm has no output.

5.2 Method 2

The method was applied on the agar dataset. Figure 15 and 16 shows the input bands used, and the reconstructed images for the 3 and 5 mm case, respectively.

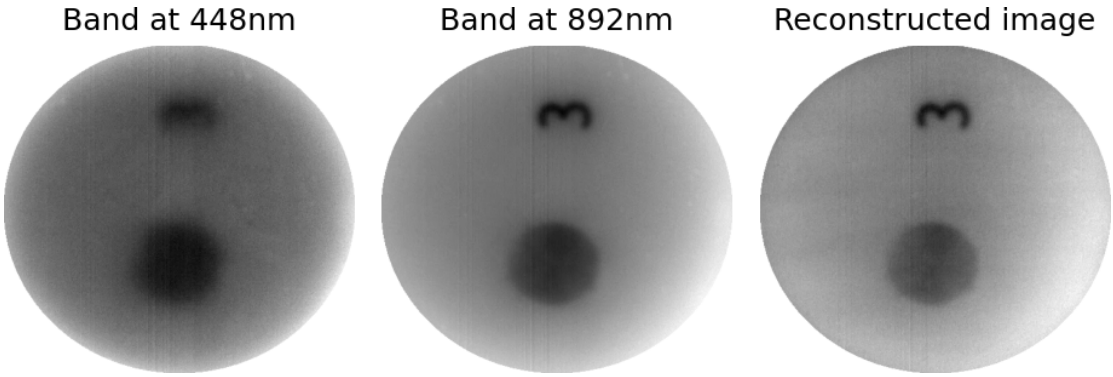


Figure 15: Inputs and outputs of method 2 when applied on a hyperspectral image of a structure covered by 3 mm turbid media. Left and middle: Bands from the hyperspectral image used as inputs to the method. Right: The output of the method.

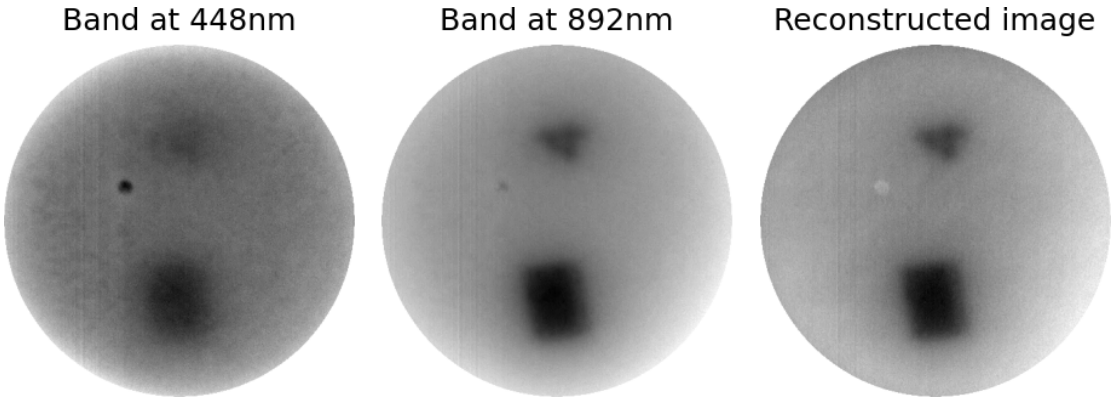
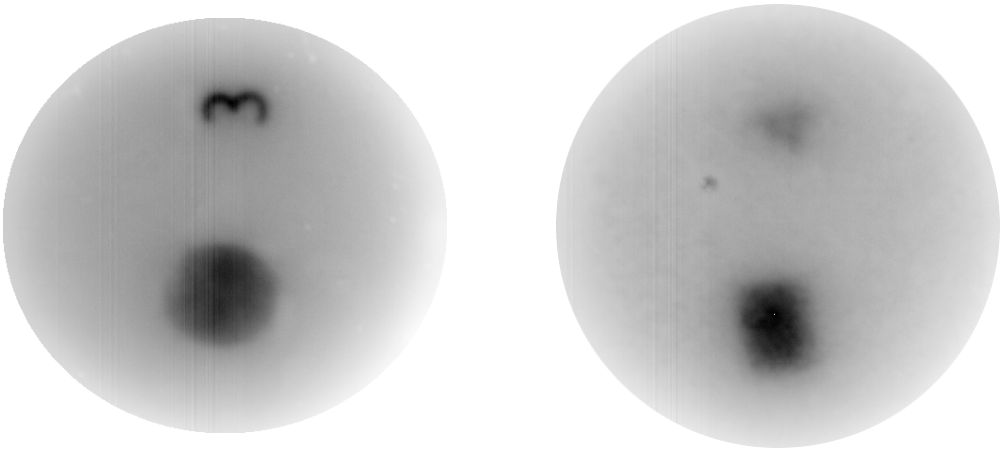


Figure 16: Inputs and outputs of method 2 when applied on a hyperspectral image of a structure covered by 5 mm turbid media. Left and middle: Bands from the hyperspectral image used as inputs to the method. Right: The output of the method.

5.3 Method 3

The method was applied on the two of the images in the agar dataset with 3 and 5 mm agar covering the structures. Figure 17a and 17b shows the reconstructed images for the 3 and 5 mm case, respectively.



(a) (b)

Figure 17: Outputs from method 3 when applied on hyperspectral images of a structures covered by 3 mm (a) and 5 mm (b) turbid media.

6 Evaluation

The three methods explored in this thesis was evaluated by comparing the results with the ground truth shown in figure 6. There are many ways to do this. The methodology used is presented in section 6.1 and the results of the evaluation is presented in section 6.2.

6.1 Methodology

The outputs of the three methods, the raw data images and the ground truth image was all normalized after masking, such that black in the image represents 0 while white represents 1.

All three methods have the same goal: To remove the effects of scattering from the input images. The ground truth in figure 6 shows what the result of this should look like. Therefore, we can compare the outputs from each method with the ground truth to evaluate how well the goal was achieved. A natural way to do this would be to compare every point of the structure in the method outputs to the corresponding point of this structure in the ground truth image. Since the structures in the images are aligned differently, this would require re-alignment if the structures. Doing this accurately without introducing pixel artifacts as a result of image rotation is very computationally demanding. Additionally, a small error in the alignment would produce inaccurate comparisons. Therefore, a simpler method to compare the two structures is used. A line is drawn through one of the structures. The corresponding line in the other structure is found manually using the image software Paint.net. Along each line, the pixel values of the image is sampled at 1000 points using linear interpolation. The two resulting pixel profiles can be compared. The pixel profiles capture the profile of the structure's edges well. This makes them suitable for use in evaluation, as it is the edge profile that is most affected by scattering.

There are multiple ways of measuring the similarity or dissimilarity between the pixel profiles. The most straightforward way is using mean square error (MSE). A perfect reconstruction would have zero MSE. However, reconstructions that look good visually may perform poorly according to this metric. If a reconstructed profile has identical shape to the ground truth profile, but is offset by a constant value, the MSE is large. Therefore, normalized correlation between the profiles is also used in the evaluation. It is a measure of similarity that is less affected by constant offsets and linear scaling than MSE. In this context we define the normalized correlation between two signals x and y by

$$corr(x, y) = \frac{\sum_{i \in \mathbb{N}} (x[i] \cdot y[i])}{\sqrt{\sum_{i \in \mathbb{N}} x[i]^2} \cdot \sqrt{\sum_{i \in \mathbb{N}} y[i]^2}} \tag{47}$$

6.2 Evaluation results

Sampled curves for all methods, the raw data, and the ground truth for the structure covered by 3 mm and 5 mm turbid media is shown in figure 18 and 19, respectively. MSE and normalized correlation values for all profiles when compared to the ground truth profile are shown in table 4.

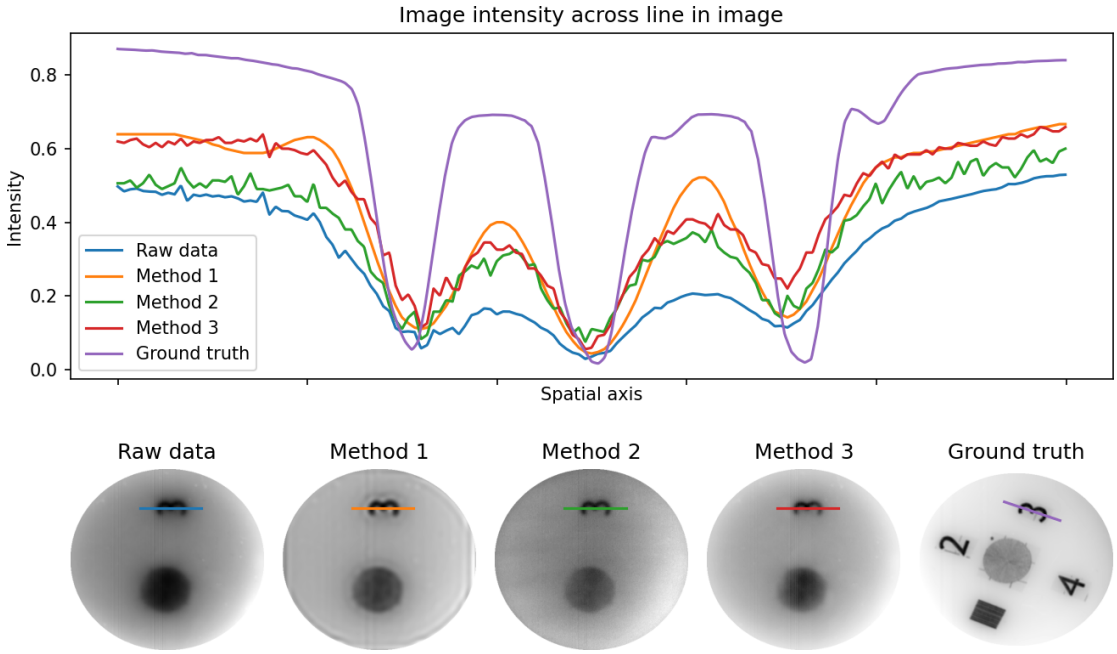


Figure 18: Top: Pixel values of the images below, sampled along the drawn line. Bottom, from left to right: The raw data image, the outputs of the three methods, and the ground truth image, all for the 3 mm case.

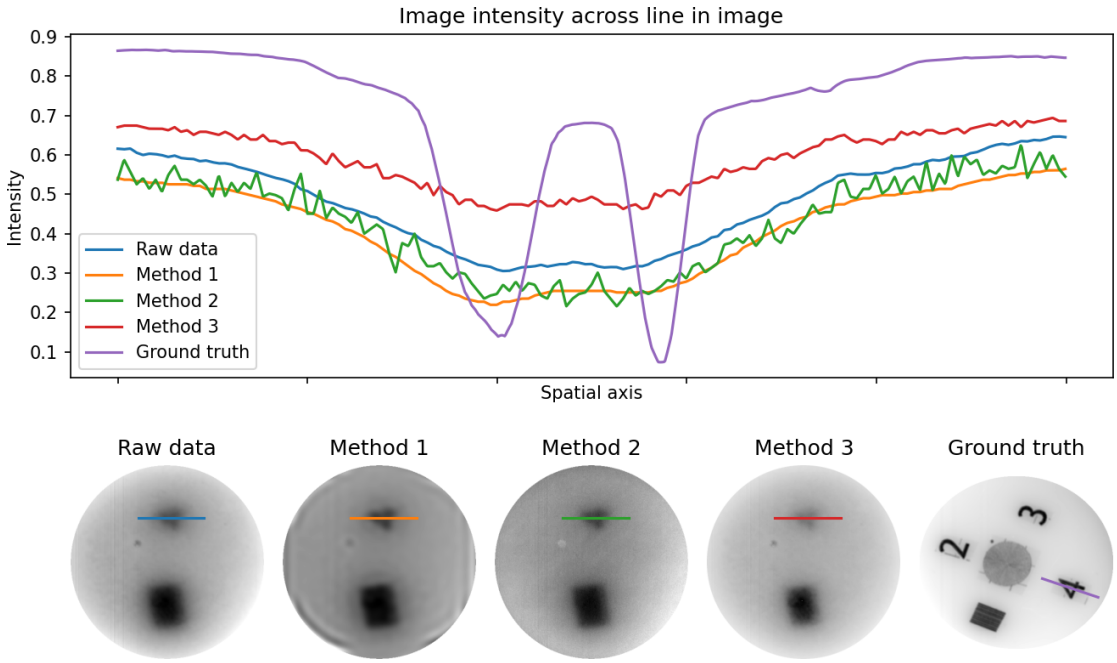


Figure 19: Top: Pixel values of the images below, sampled along the drawn line. Bottom, from left to right: The raw data image, the outputs of the three methods, and the ground truth image, all for the 5 mm case.

Table 4: MSE and normalized correlation values for all profiles when compared to the ground truth profile. Lower MSE is better. Higher normalized correlation is better.

(a) 3 mm case

Pixel profile	**MSE**	**Normalized correlation**
Raw data	.1354	.9560
Method 1	.0497	.9848
Method 2	.0857	.9820
Method 3	.0541	.9790
Ground truth	0	1

(b) 5 mm case

Pixel profile	**MSE**	**Normalized correlation**
Raw data	.0694	.9828
Method 1	.1037	.9845
Method 2	.0941	.9823
Method 3	.0386	.9771
Ground truth	0	1

7 Discussion

A discussion of the results of the three methods and their evaluation is given in section 7.1-7.3. Section 7.4 contains a discussion on the evaluation methodology.

7.1 Method 1

Looking at figure 13, it is evident that the deconvolition algorithm has worked as intended. When deconvolving with the PSF at increasing depths, the resulting image initially becomes sharper, with more defined edges. Then, after a certain depth, deconvolution artifacts become visible, and increasingly prominent with increasing depths. Determining the optimal depth for deconvolution becomes a trade-off between sharp edges and deconvolution artifacts. Quantifying these phenomena accurately is challenging, making finding the optimal depth challenging. 12 different configurations of the algorithm where tested on two different images. Only one of the 24 cases resulted in a prediction for the optimal depth, which is alarmingly low. There are two possible explanations for this. The fist one is that the combination of denoising the inputs, Richardson-Lucy deconvolution, and the ENR focus measure is the only configuration that can reliably determine the optimal depth, but only up to a certain level of blurriness of the input image. In this case, method 1's fail to produce an output for the 5 mm case can be attributed to that the input was too blurry to reconstruct. The second possible explanation is that the method is unstable, and that the peak in the focus curve observed for the one configuration is random. One can imagine that for a different dataset, another configuration might produce a peak in the focus curve. The fact that the working configuration is the same as the one used by Jiang et. al. [4] for CT images strengthens the first explanation. To test the reliability of the focus finding part of this method, a larger dataset is needed.

One might imagine that if the depth is known or determine by other means, then deconvolving at that depth would be a reliable deblurring method. In theory, this is true. If the model for the PSF is accurate, and all PSF parameters are known, then deconvolution should produce an accurate reconstruction. However, by fixing all parameters, the method looses one of it's strengths: By leaving the depth as a degree of freedom for the algorithm to determine, this parameter can be used to compensate for inaccuracies in other parameters of the PSF, such as the scattering coefficient. This is because both the depth and scattering coefficient affect the most important characteristic of the PSF: Its width. This benefit requires an accurate and stable focus finding component of the algorithm.

The output of the method for the 3 mm case has sharp edges, and includes some signs of deconvolution artifacts. This is consistent with the result Jiang et. al. got in [4].

7.2 Method 2

In the 3 mm case, figure 17a shows that visually, the method provided a small but noticeable improvement. This improvement is confirmed by a lower MSE and higher normalized correlation than the raw data when compared to the ground truth. Visually, these results seems to be similar to the ones Miura and Sato got in [9]. Miura and Sato does not provide any numerical quality improvement metrics that allow for quantitative comparison of the results.

In the 5 mm case, it is hard to argue that the results shown in 17b are better than the input, and that the method provided an increase in image quality. This is backed by that the output has higher MSE and lower normalized correlation than the raw data when compared to the ground truth.

7.3 Method 3

In the 3 mm case, comparing figure 17a and 7 shows that visually, the method provided a small but noticeable improvement. This improvement is confirmed by a lower MSE and higher normalized correlation than the raw data when compared to the ground truth. Method 3 differentiates itself from the two other methods in that it does not operate spatially, but only spectraly. It may seem counter-intuitive that this method is able to improve sharpness of edges, as this is a spatial phenomenon. An possible explanation of why the method works is given below.

Different channels are blurred in different degrees, based on the scattering coefficient of the band's wavelength. This means that the structure's spectrum "leaks" into the surrounding medium differently at different wavelengths. Thus, the spectrum at the edge looks substantially different than the spectrum of the structure, creating a improved sharpness in the image displaying the correlation of each pixels spectrum and the ground truth spectrum.

Method 3 is also different from the other methods in that it uses data from the ground truth image. One can see this as giving the method an unfair advantage. However, in most real world application, the absorbance spectra of the constituents of an examined object are well documented, so this is not seen as an issue.

In the 5 mm case, method 3 performs better than the raw data when using MSE as a performance metric. Despite this, it is hard to argue that the method provided any improvement in quality that was visible to the human eye. When using normalized correlation as a performance metric, the method performed worse than the raw data.

7.4 Evaluation of deblurring

The goal of all three methods is to reduce blurring created by scattering of light in turbid media. Objectively describing what the perfect reconstruction would look like is trivial: It should look like the ground truth image. Objectively quantifying how good an imperfect reconstruction is, however, is challenging. Which reconstruction is "best" out of two imperfect ones, depends entirely on the way performance is quantified.

In this thesis the three methods were evaluated based on two numerical performance measures: The MSE- and the normalized correlation between pixel profiles in the reconstructed images and ground truth image. The images shown in figures 13, 14, 15, 16, 17a and 17b can also be used to perform visual quality assessment.

In the 5 mm case, when using MSE as the evaluation metric, method 3 greatly out-preforms the two other, even though visually, this does not seem to be the case. In the 3 mm case, when using normalized correlation as the evaluation metric, method 1 performs best, despite that the artifacts resulting from method 1 make in less visually pleasing than the result from the other methods. This shows that MSE, normalized correlation and visual inspection provide different views on what is "good". MSE is an error measure, it weighs every sample along the pixel profile equally, and is greatly affected by constant and linear scaling. Normalized correlation is much more concerned with the general shape of the profiles, and is less affected by constant and linear scaling. Visual inspection is in itself a subjective process, depending on both the individual and the viewing circumstances. However, most people share some common perceptual preferences, with the most important factor being how recognizable objects of an image are.

Because of the difficulty in objectively evaluating performance, it is unreasonable to pick a "best" method. In a practical application, where the use of the reconstructed images is known, it is likely easier to make this decision.

8 Conclusion

This thesis has explored, developed and evaluated methods for processing hyperspectral images that reduce blurring created by scattering of light in turbid media. All three implemented methods measurably and noticeably improved the quality of the least blurry image in the dataset. When applied on the most blurry image, all three methods performed poorly.

The main contribution of this thesis is the adaptation of the three methods and the evaluation of their performance when applied on a hyperspectral dataset. This thesis has shown that the focus-finding principles developed by Jiang et. al.[4] works in combination with the PSF derived by Shimizu et. al [11] in the context of hyperspectral images. It has been shown that the scattering correction methods by Svaasand et. al.[15] and Bjorgan et. al. [1] are useful for spatial deblurring. Finally, the thesis has compared the performance of these methods numerically and visually.

References

[1] Asgeir Bjorgan, Matija Milanic, and Lise Lyngsnes Randeberg. Estimation of skin optical parameters for real-time hyperspectral imaging applications. *Journal of biomedical optics*, 19(6):066003, 2014.

[2] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on image processing*, 16(8):2080–2095, 2007.

[3] Richard C Haskell, Lars O Svaasand, Tsong-Tseh Tsay, Ti-Chen Feng, Matthew S McAdams, and Bruce J Tromberg. Boundary conditions for the diffusion equation in radiative transfer. *JOSA A*, 11(10):2727–2741, 1994.

[4] Ming Jiang, Ge Wang, Margaret W Skinner, Jay T Rubinstein, and Michael W Vannier. Blind deblurring of spiral ct images. *IEEE Transactions on medical imaging*, 22(7):837–845, 2003.

[5] Remi Johansen. Konstruksjon og avbildning av hudfantomer av agargel. Project report in TFE4530, Department of Electronic Systems, NTNU – Norwegian University of Science and Technology, Jan. 2008.

[6] Eui Chul Lee and Kang Ryoung Park. Image restoration of skin scattering and optical blurring for finger vein recognition. *Optics and Lasers in Engineering*, 49(7):816–828, 2011.

[7] Anat Levin, Yair Weiss, Fredo Durand, and William T Freeman. Understanding and evaluating blind deconvolution algorithms. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1964–1971. IEEE, 2009.

[8] Hashim Mir, Peter Xu, and Peter Van Beek. An extensive empirical evaluation of focus measures for digital photography. In *Digital Photography X*, volume 9023, pages 167–177. SPIE, 2014.

[9] Naoto Miura and Yoichi Sato. Deblurring vein images and removing skin wrinkle patterns by using tri-band illumination. In *Asian Conference on Computer Vision*, pages 336–349. Springer, 2012.

[10] Lawrence A Shepp and Yehuda Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE transactions on medical imaging*, 1(2):113–122, 1982.

[11] Koichi Shimizu, Koji Tochio, and Yuji Kato. Improvement of transcutaneous fluorescent images with a depth-dependent point-spread function. *Applied optics*, 44(11):2154–2161, 2005.

[12] Thorsten Spott. *Characterization of layered tissue structures with diffusely propagating photon-density waves*. Norwegian University of Science and Technology, Faculty of Electrical Engineering and Telecommunications, Department of Physical Electronics, Trondheim, 1999.

[13] Libin Sun, Sunghyun Cho, Jue Wang, and James Hays. Good image priors for non-blind deconvolution. In *European conference on computer vision*, pages 231–246. Springer, 2014.

[14] Yu Sun, Stefan Duthaler, and Bradley J Nelson. Autofocusing in computer microscopy: selecting the optimal focus algorithm. *Microscopy research and technique*, 65(3):139–149, 2004.

[15] Lars O Svaasand, LT Norvang, EJ Fiskerstrand, EKS Stopps, MW Berns, and JS Nelson. Tissue parameters determining the visual appearance of normal skin and port-wine stains. *Lasers in Medical Science*, 10(1):55–65, 1995.

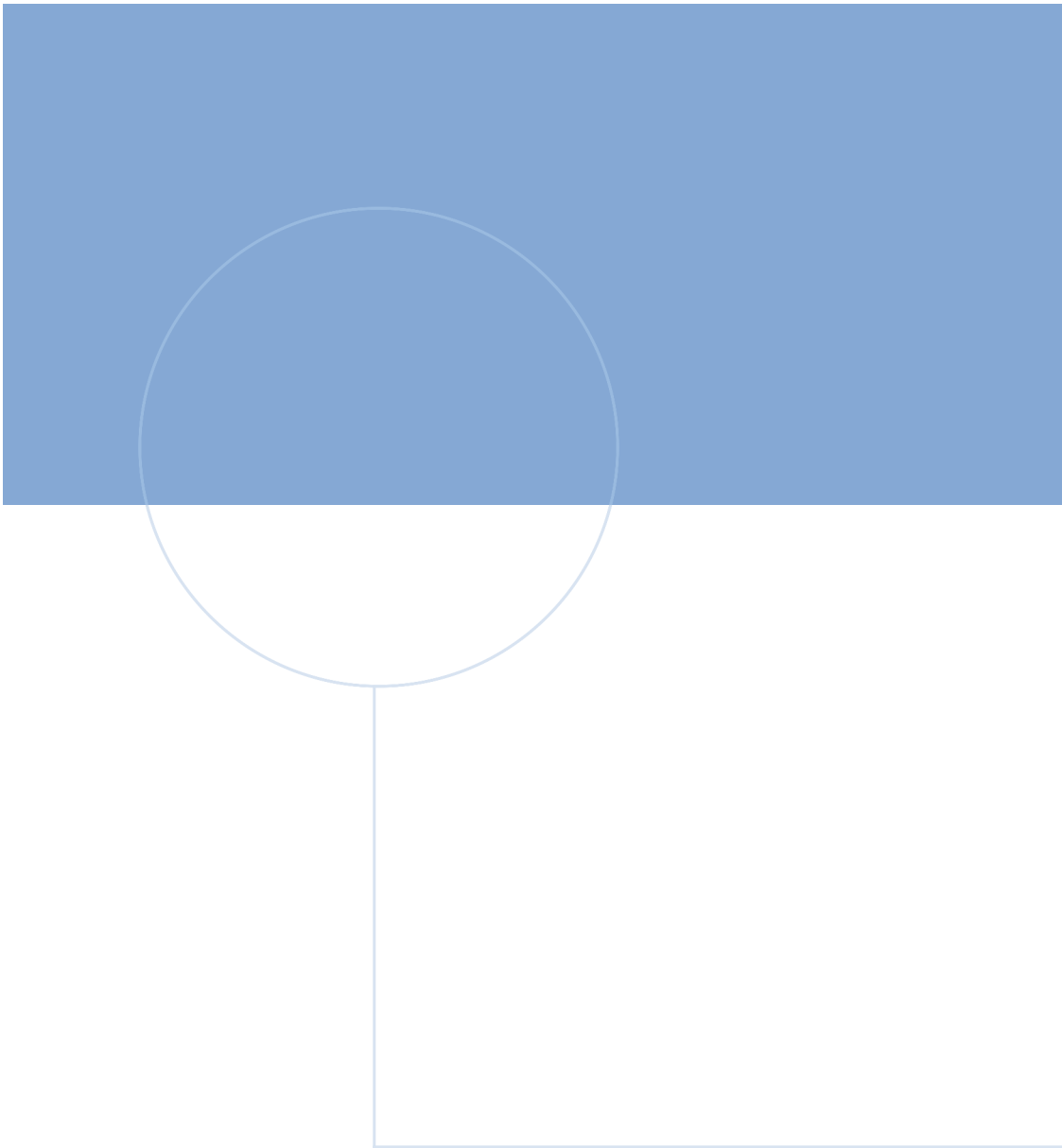
[16] Torstein Tenstad. Scattering correction methods. Project report in TFE4580, Department of Electronic Systems, NTNU – Norwegian University of Science and Technology, Dec. 2021.

[17] Hugo J Van Staveren, Christian JM Moes, Jan van Marie, Scott A Prahl, and Martin JC Van Gemert. Light scattering in lntralipid-10% in the wavelength range of 400–1100 nm. *Applied optics*, 30(31):4507–4514, 1991.

[18] Wikipedia contributors. Wiener deconvolution — Wikipedia, the free encyclopedia, 2022. [Online; accessed 9-June-2022].

[19] Huangxuan Zhao, Ziwen Ke, Ningbo Chen, Songjian Wang, Ke Li, Lidai Wang, Xiaojing Gong, Wei Zheng, Liang Song, Zhicheng Liu, et al. A new deep learning method for image deblurring in optical microscopic systems. *Journal of biophotonics*, 13(3):e201960147, 2020.

[20] Caigang Zhu and Quan Liu. Review of monte carlo modeling of light transport in tissues. *Journal of biomedical optics*, 18(5):050902, 2013.



NTNU

Norwegian University of
Science and Technology